Vanishing Point Detection: Representation Analysis and New Approaches

Virginio Cantoni, Luca Lombardi, Marco Porta, Nicolas Sicard
Dipartimento di Informatica e Sistemistica – Università di Pavia
Via Ferrata, 1 – 27100 – Pavia – Italy
{cantoni, luca, porta, sicard}@vision.unipv.it

Abstract

In this paper we introduce two different representation approaches and propose two techniques to estimate the position of vanishing points in an image, one based on a probabilistic strategy and the other focused on a deterministic analysis. Unlike most of the methods so far developed, which exploit the Gaussian sphere, the new techniques operate in the (ρ, θ) polar parameter space and in the (x, y) image plane coordinate space. Both the solutions will be described and compared, also through the discussion of the results obtained from their application to real images.

1. Introduction

As one can easily note, most real images show some kind of perspective distortion. Given a set of parallel lines in the three-dimensional (3D) space, they will converge to a single point on the image plane, called *vanishing point*. The knowledge of vanishing points is an important step toward 3D interpretation, allowing meaningful information to be obtained about the real scene (such as depth, object dimensions, etc.). Although this approach may suffer from several problems due to the irregularity of natural settings, it turns out to be especially useful for man-made environments, where many regular structures and parallel lines are present.

Essentially, it is possible to distinguish three kinds of approaches for finding vanishing points in an image. Even though they all are based on the concept of Hough Transform [1] [2], they differ in the parameter space where votes are accumulated. Most of the methods so far developed exploit the Gaussian sphere to represent orientations, i.e. they operate in the space of spherical coordinates (θ, φ). In this paper, we will describe two alternative techniques which work, respectively, in the (ρ, θ) polar parameter space and in the (x, y) image plane coordinate space.

The paper is structured as follows. Section 2 briefly describes the main characteristics of methods based on the Gaussian sphere. Section 3 introduces the bases for approaches in the polar space. Sections 4 and 5 discuss our proposals in detail. Section 6 presents experimental results. Section 7, at last, draws some conclusions and gives hints for future work.

2. Gaussian-sphere-based approaches

Techniques based on the Gaussian sphere, first introduced by Barnard [3], use a unit radius sphere centered in the optical center as an accumulation space. For each line segment detected on the image plane, the *interpretation plane* identified by the segment and by the optical center is considered. This plane will intersect the Gaussian Sphere to form a great circle, i.e. a circumference on which votes are accumulated. Since circumferences resulting from line segments representing parallel lines in the real scene intersect each other, the points on the sphere which, at the end of the accumulation process, have more votes, can be considered as those corresponding to vanishing points' directions. On the image plane, such points are then identified by finding the intersections of direction straight lines with the image plane itself.

Barnard's method has the main advantage of being able to represent both the finite and infinite vanishing points. However, it has also several limitations. Since a regular quantization in θ and φ does not generate cells of equal surface, the count of votes on the Gaussian sphere must rely on some form of statistical approach which considers the different areas. To solve this problem, in [4], for example, the sphere is sampled from a coarse resolution to a fine resolution, using a hierarchical Hough Transform. Another possible solution, proposed in [5], is to use a quantization in "semiregular" rectangular cells, based on a regular quantization in φ and an irregular quantization in θ.

In general, however, for the results obtained through the Barnard's method to be reliable, it is necessary that orientations of interest in the 3D space correspond to main groups of converging line segments in the image [6], and this depends on the viewing angle and on the distance to the scene. Moreover, textural effects caused by natural patterns and artifacts of digital image geometry can combine to produce spurious maxima (i.e. a sort of noise).

Since lines assumed parallel in the 3D space have their images only approximately concurrent in the image plane, techniques have been developed to model the orientation error [6], as well as fully statistical approaches for managing uncertainty (such as those described in [7] and [8]). Improvements can also be obtained when the geometry of objects of interest in the scene is known, by exploiting *primitive-based* techniques [6]. In such case, the search for vanishing points is carried out by imposing constraints on objects shapes, which can be easily translated into orientation constraints on the Gaussian sphere.

3. Approaches in the polar space

In a polar parameter space, points in the image plane are mapped to sinusoids, according to the following equation:

$$\rho = x \cos\theta + y sin\theta \qquad (1)$$

where x and y are the coordinates of a point P in the image plane, while ρ and θ describe the straight line passing through P and oriented according to the phase of the local gradient (Figure 1).
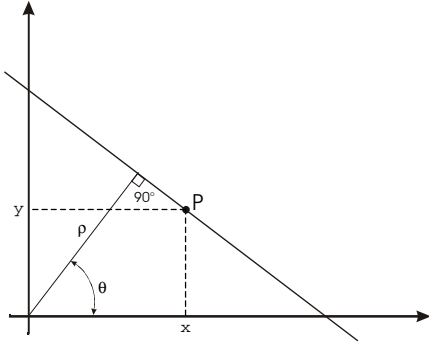

Figure 1. Characterization of a straight line through polar parameters

Several methods use an approach based on the Hough Transform accumulation space in the polar plane. Nakatami [9], for example, determines the straight lines on the image plane and then uses a deterministic approach to find the points of convergence. Precision problems arising in case of thick edges, however, limit its practical applicability.

The main advantage of using a polar space is that parameters ρ and θ are limited (for an LxL image, ρ is within the range [0, √2L), while θ takes values in the range [0,2π[). Points belonging to the same straight line in the image plane have corresponding sinusoids which intersect in a single point in the polar space. Since in a vanishing point several straight lines converge, its sine curve will be given by the sum of all the sinusoids of the points

on the lines. This characteristic will be exploited in the method we are going to describe in the next section.

4. First proposal

This technique [10] uses a statistical approach to search for the sine curve corresponding to the vanishing point in the polar parameter space. Essentially, the pair of parameters (x_0, y_0) representing the point is estimated through a least square method.

The approach consists in a minimization of the following functional:

$$\min_{x_0, y_0} \sum_{i=1}^{n} W_i (\rho_i - x_0 \cos\theta_i - y_0 sin\theta_i)^2 \qquad (2)$$

where:

$$W_i = \frac{v_i}{V} \qquad (3)$$

In expression (3), v_i is the number of times that a pair (ρ_i, θ_i) is observed, while V is the total number of votes (points) in the polar diagram. Practically, expression (2) does nothing but searching for that point $P_0(x_0, y_0)$ in the image plane which minimizes the distance from all the straight lines observed on it.

Setting $a_i = cos\theta_i$, $b_i = sin\theta_i$ and deriving with respect to x_0 and y_0, we get the following couple of equations:

$$\begin{cases} \sum_{i=1}^{n} W_i a_i (\rho_i - a_i x_{0i} - b_i y_0) = 0 \\ \sum_{i=1}^{n} W_i b_i (\rho_i - a_i x_{0i} - b_i y_0) = 0 \end{cases} \qquad (4)$$

Then, we can call:

$$A = \sum_{i=1}^{n} W_i a_i^2 \ ; \ B = \sum_{i=1}^{n} W_i b_i^2 \ ; \ C = \sum_{i=1}^{n} W_i a_i b_i$$

$$D = \sum_{i=1}^{n} W_i a_i \rho_i \ ; \ E = \sum_{i=1}^{n} W_i b_i \rho_i \qquad (5)$$

Since A, B, C, D and E are constants, x_0 and y_0 can be simply found by solving the following linear system:

$$\begin{cases} Ax_0 + Cy_0 = D \\ Cx_0 + By_0 = E \end{cases} \qquad (6)$$

which is straightforward.

After the first estimation of the vanishing point, the process is repeated by suppressing the outliers. Letting:

$$\varepsilon_i = \rho_i - \overline{\rho}_i \qquad (7)$$

be the residual error, where ρ_i is the value of the ρ pa-

rameter for the ith point in the polar plane (ρ, θ) and $\overline{\rho}_i$ is the corresponding value on the sine curve representing the vanishing point, the variance of the residuals is:

$$\sigma^2 = \sum_{i=1}^{n} W_i \varepsilon_i^2 \qquad (8)$$

All those points satisfying the following condition are then considered outliers:

$$\left|\rho_i - \overline{\rho}_i\right| > k\sigma \qquad (9)$$

where k is generally between 2 and 3.

The analysis just described is applied recursively until the Marshall distance between two successive estimations of the vanishing point is less than the desired precision (i.e., according to disequation (9), until no more outliers can be found). Figure 2 shows the polar space before (a) and after (b) the outliers elimination process has been carried out in an example image.
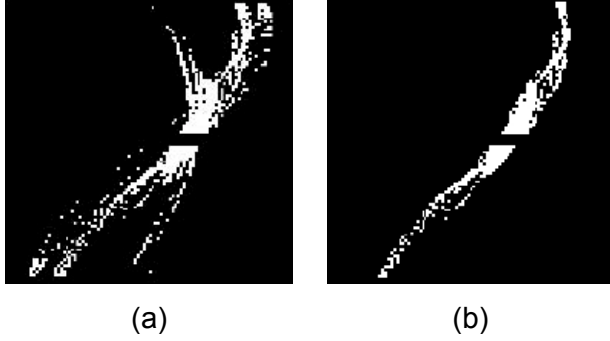

(a)   (b)

Figure 2. (a) Original parameter space containing outliers; (b) Parameter space after the outliers removal

5. Second proposal

This second technique works directly in the image plane, which is used as an accumulation space for a particular version of the Hough Transform.

The algorithm is composed of the following steps. First, an edge detection operation is performed on the original image through an isotropic operator, which is composed of the following two 3x3 masks (one for the horizontal and one for the vertical component):

$$D_x = \begin{bmatrix} -1 & 0 & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ -1 & 0 & 1 \end{bmatrix} \quad D_y = \begin{bmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & -\sqrt{2} & -1 \end{bmatrix}$$

The result image is then normalized and thresholded against a very high value, to eliminate redundant informa-

tion. This way, a binary image is obtained. For each point P within it, the tangent is calculated, by means of the following expression:

$$m = tg\,\theta = \frac{\left|D_y\right|}{\left|D_x\right|} \qquad (10)$$

where D_x and D_y are the x and y components resulting from the application of the isotropic operator in the point. Now that θ is known, it can be used to draw a straight line, with slope m, passing through P. Such line is then accumulated in the (x, y) parameter space. At the end of the process, those points having the greatest numbers of votes will be marked as candidates.

Images of both indoor and outdoor scenes are often full of (nearly) vertical and horizontal lines. Since, for the most part, they would identify vanishing points at the infinite (which are not useful for our investigation), we introduce the following constraint:

$$\frac{1}{p} < \left|tg\,\theta\right| < q \qquad (11)$$

where p and q have values greater than one and are chosen according to the particular needs (for example, $p = q = 64$). Without the limitation imposed by condition (11), false detections would result from intersections between lines actually converging to vanishing points and (nearly) horizontal or vertical lines.

Expression (11), which has the implicit advantage of limiting the values assumed by parameter $tg\,\theta$, is based on the assumption that the camera is perpendicular to the acquisition plane and horizontal lines are parallel to the x axis, as occurs in many real images. If such assumption fails, a different angle should be taken into account.

In order to improve the precision of maxima detection in the (x, y) parameter space, a cumulative procedure is applied by averaging the values contained in an 11x11 window around each pixel. Moreover, to improve their placement, vanishing points are considered in the barycenters of the pixels contained in neighborhoods of the maxima locations.

Of course, a vanishing point may also lie outside the image window. In this case, the parameter space must be extended, so that points whose coordinates are greater than the image size can be represented, and the areas of neighborhoods used in the smoothing process must be enlarged.

6. Experimental results

In this section we will present some experimental results obtained from the application of the proposed methods to real images. Such results, however, should not be considered per se, but rather as starting points for more

elaborate strategies able to combine the two approaches, to exploit the best from each. For example, the vanishing point search in the (x, y) parameter space could be used as a first stage to select "candidate" straight lines in the image, which could then be considered by the technique operating in the polar space.

Figure 4 shows the straight lines obtained from the application of the Hough Transform in the (ρ, θ) polar space to the image in Figure 3. Figure 5, instead, shows the accumulation parameter space obtained with the method operating in the (x, y) coordinate system.

In Figures 6, 7 and 8 there are analogous results obtained from a different image.

The coordinates of the vanishing points found through the first and the second method for the image in Figure 3 are, respectively, (305, 198) and (306, 200).

Although in this case the first estimation is better than the second, for images in which more than one vanishing point can be identified or where more complex patterns are present (such as in Figure 6), the second technique turns out to be more reliable.

The analysis proposed here, however, is only qualitative. We will soon carry out more precise investigations to combine and quantitatively compare our techniques, both with each other and with the approaches described in Section 2.

7. Conclusions

In this paper we have summarized the representation approaches for estimating vanishing points. Two new techniques have been presented, which differ in the parameter space used to accumulate votes through the Hough Transform. While the first method, based on a probabilistic approach, operates in the (ρ, θ) polar space, the second, focused on a more deterministic analysis, works directly in the image plane (x, y) coordinate space. This second technique is computationally lighter than the first, since it does not require to iteratively solve a linear system, nor to compute variance values. Nevertheless, it often produces results of higher precision. It is also interesting to note that the algorithms described could be used to detect the focus of expansion/contraction in motion analysis (for example, through the optical flow).

Generally, in many real images, more than one vanishing point is present. In these cases, our proposals find the vanishing point produced by the greater number of straight lines in the image. However, other vanishing points can be found in successive analyses by excluding the straight lines covered by the points already detected.

The two approaches described could also be combined, to get more reliable results. For example, the second technique could be used to confirm the vanishing points statistically found through the first one.

Figure 3. Original image for the first experiment

Figure 4. Straight lines found through the Hough Transform in the polar space applied to the image in Figure 3
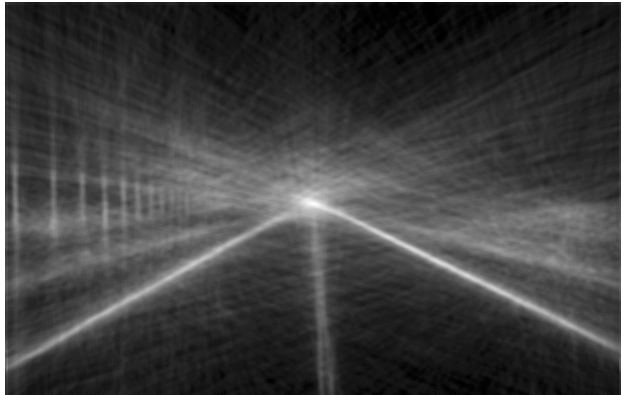

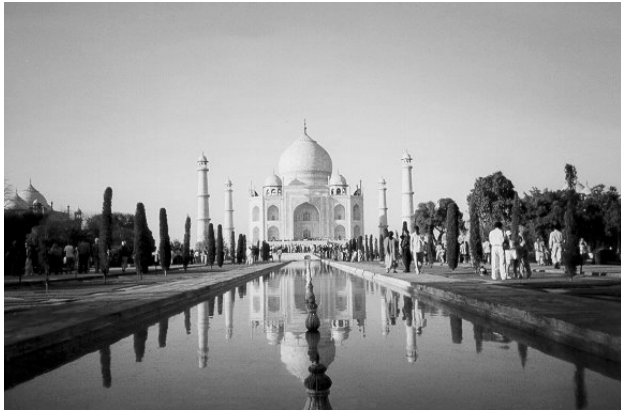**Figure 5. Accumulation of votes in the (x, y) parameter space for the image in Figure 3**

Figure 6. Original image for the second experiment

Figure 7. Straight lines found through the Hough Transform in the polar space applied to the image in Figure 6
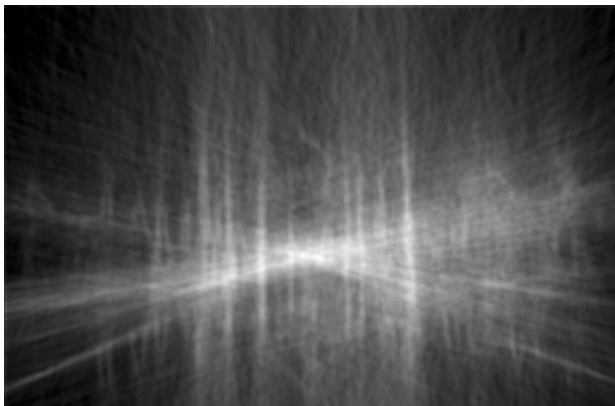

Figure 8. Accumulation of votes in the (x, y) parameter space for the image in Figure 6

Or also, the first approach could be modified to take into account only straight lines selected through the second strategy. Our work in the immediate future will be devoted to such analyses, as well as to a quantitative comparison of the proposed methods with those based on the Gaussian sphere.

8. References

[1] Hough, P. V., "Methods and means to recognize complex patterns", *U.S. Patent 3,069,654*, 1962.

[2] Ballard, D., Brown, C., *Computer Vision*. Prentice Hall, Englewood Cliffs, NJ, 1982.

[3] Barnard, S. T., "Interpreting Perspective Images ". *Artificial Intelligence*, vol. 21, pp. 435-462, 1983.

[4] Quan, L., Mohr, R., "Determining perspective structures using hierarchical Hough transform". *Pattern Recognition Letters*, 9, pp. 279-286, 1989.

[5] Lutton, E., Maitre, H., and Lopez-Krahe, J., "Contribution to the Determination of Vanishing Points Using Hough Transform". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 4, April, pp. 430-438, 1994.

[6] Shufelt, J. A.., "Performance Evaluation and Analysis of Vanishing Point Detection Techniques". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 3, March, pp. 282-288, 1999.

[7] Collins, R. T., Weiss, R. S., "Vanishing Point Calculation as a Statistical Inference on the Unit Sphere". *Proceedings of the Third International Conference on Computer Vision*, pp. 400-403, 1990.

[8] Brillault-O'Mahony, B., "New Method for Vanishing Point Detection". *Computer Vision, Graphics, and Image Processing: Image Understanding*, vol. 54, no. 2, pp. 289-300, 1991.

[9] Nakatami, H., Kimura, S., and Saito, O, "Extraction of vanishing point and its application to scene analysis based on image sequence". *Proceedings of the 5th International Conference on Pattern Recognition*, pp. 370-372, 1980.

[10] Matessi, A., Lombardi, L., "Vanishing Point Detection in the Hough Transform Space". *Proceedings of the Fifth International Euro-Par Conference*, Tolouse, France, August - September, pp. 987-994, 1999.